<u>VIA ECF</u>

June 10, 2025
Hon. Ona T. Wang
Southern District of New York

   *In re OpenAI, Inc.*, *Copyright Infringement Litigation*, 25-md-3143 (SHS) (OTW)
   This Document Relates To: No. 1:23-cv-11195 (SHS) (OTW)

Dear Magistrate Judge Wang:

   I write on behalf of News Plaintiffs[1] to request an order compelling Defendants to produce documentation concerning their metric-tracking dashboards, which will facilitate News Plaintiffs' efforts to obtain relevant data about how news content (including that of the News Plaintiffs) is delivered through Defendants' Generative AI models, products, and services.[2] This discovery is relevant to News Plaintiffs' allegations that Defendants' products divert traffic away from their sites, as well as to revealing the extent to which Defendants' models and products exploit, reference, and infringe copyrighted works. Ultimately, News Plaintiffs are asking Defendants to produce data reflecting metrics such as the number of prompts or queries a user inputs per session, and the extent to which users "click through" to publisher sites (as opposed to remaining within Defendants' platforms). But as a first step to obtaining that data, News Plaintiffs seek documentation concerning various metric-tracking dashboards identified by News Plaintiffs during discovery. Once News Plaintiffs receive this documentation, the parties can work together on a plan for Defendants to query the dashboards and produce responsive data.

## I. Defendants' Metric-Tracking Dashboards are Highly Relevant and Contain Responsive Information to News Plaintiffs' Requests

   News Plaintiffs seek to examine Defendants' metric-tracking dashboards for their generative AI models, products, and services to collect metrics on, *inter alia*, clickthrough rates, fallback rates, bounce rates, chat impressions, and paywall violations.

   Neither OpenAI nor Microsoft dispute that their dashboards and their associated metrics are relevant to these cases – to the contrary, News Plaintiffs have sought these metrics since the beginning of these cases, and Defendants agreed to produce them. *See Times* Dkt. 285-1 at 13 (RFP No. 39 seeking "[d]ocuments concerning the study, testing, or evaluation of Defendants' Generative AI Products and Services and Generative AI Models for hallucinations, memorization, and any problematic behaviors…."); *Times* Dkt. 379-5 at 3 (RFP No. 47 seeking "[d]ocuments concerning user engagement metrics, events, and analytics data collected during each user session, including duration, the number of prompts or queries that a user inputs per session, click-through

---

[1] This motion is filed on behalf of News Plaintiffs The New York Times Company ("The Times"), the *Daily News* Plaintiffs, the Center for Investigative Reporting, The Intercept, and the *Ziff Davis* Plaintiffs.

[2] Microsoft has agreed to produce some of the requested documentation and information, as explained below in Section III, *infra*. If Microsoft fails to produce that documentation promptly, and fails to begin querying and producing responsive data from its dashboards, News Plaintiffs intend to raise this issue against Microsoft at the June 25, 2025 hearing. News Plaintiffs have not reached any agreement with OpenAI.

data, fallback rates, and bounce rates.").[3] Yet now that News Plaintiffs have discovered the dashboards where this data is kept, Defendants inexplicably refuse to produce the data or explain how the dashboards can be queried to obtain the data.

## II.    Defendants Are Refusing to Produce Responsive Metrics

Both OpenAI and Microsoft maintain extensive metrics about how users engage with their generative AI products, models, and services. For example, when asked "[w]hat sort of running metrics" OpenAI keeps on user outputs, Michael Trinh, OpenAI's Associate General Counsel for Litigation and Discovery, responded ███████████." Dkt. 44-7 (Transcript of the January 29, 2025 deposition of Michael Trinh) at 124 (emphasis added). Mr. Trinh confirmed that OpenAI collects metrics on "████████████████████████████████" and "███████████████" with various metric-tracking dashboard tools. *See* Dkt. 44-7 at 126-32. As another example, during the damages trial in *United States v. Google,* Case No. 20-cv-3010 (D.D.C.), Nick Turley, Head of Product for ChatGPT, described statistics that track how users interact with OpenAI's products, and confirmed that "**local news**" is a "**pretty common quer[y]**" by ChatGPT users. *See* Ex. A at 396:4-397:7 (May 2, 2025) (emphasis added).

Likewise, News Plaintiffs have learned that Microsoft maintains dashboards used to track metrics responsive to News Plaintiffs' requests. *See, e.g.,* Ex. B at -60 ████████████████████████████████████████████████████████████████████████████████████████████████████; Ex. C at -52 ████████████████████████████████████████████████████████████████████████████████████████████████████"). As another example, although Microsoft stated in a July 24, 2024 blog post that it had been "look[ing] closely at how generative search impacts traffic to publishers" and represented that "[e]arly data indicates that this experience maintains the number of clicks to websites and supports a healthy web ecosystem," Microsoft has yet to produce the data to support its claims. *See* Ex. D.

Upon discovering Defendants' metric-tracking dashboards, News Plaintiffs sent discovery deficiency letters to Microsoft and OpenAI on February 10, 2025[4] (Ex. E) and February 12, 2025 (Ex. F), respectively, identifying responsive metrics tracked by these dashboards as missing from their productions and explaining how this information should have been produced as responsive to prior discovery requests. In addition, though unnecessary, on March 7 and 12 News Plaintiffs served follow-up discovery requests explicitly asking Defendants to produce documents, information, and data relating to these dashboards, where News Plaintiffs identified the relevant dashboards by name. *See* Ex. G at 10-11 (RFP Nos. 154, 156-157); Ex. H at 8-10 (RFP Nos. 143, 149 and 151). However, after months of meet-and-confer efforts, neither Defendant has produced any metrics, circularly arguing that News Plaintiffs had not identified the requested metrics with enough specificity – while refusing to disclose which metrics it tracks, and how. *See* Ex. I; Ex. J.

---

[3] The Daily News Plaintiffs served identical discovery requests on June 14, 2024.
[4] Microsoft's February, 28, 2025 response to News Plaintiffs' February 10 letter is attached as Exhibit K.

### III. News Plaintiffs Request Targeted Discovery into Defendants' Tracking Dashboards

News Plaintiffs' requests identify the requested metrics with sufficient specificity, and Defendants are familiar with these metrics given Defendants' public statements about how they track user engagement, including how the public uses their products to engage with news content. Although News Plaintiffs have been willing to work with Defendants to identify more specific metrics, Defendants have refused to provide News Plaintiffs with information about the specific functionalities and capabilities of Defendants' metric-tracking dashboards. Given the information asymmetry between News Plaintiffs and Defendants, News Plaintiffs asked both OpenAI and Microsoft to provide documentation detailing the specific metrics tracked by these dashboards and how these dashboards can be queried. OpenAI has declined to provide News Plaintiffs with any documentation or information, stalling discovery into relevant metrics. Microsoft has only agreed to provide documentation and information about "dashboards for data related to legacy consumer Bing Chat," which "cannot be queried," and is otherwise withholding documentation about dashboards that can currently be queried for the relevant metrics.[5] The end result is that both Defendants are refusing to provide information about the tools necessary for News Plaintiffs to obtain relevant metrics on Defendants' models, products, and services.

Accordingly, News Plaintiffs ask that this Court order Defendants to provide documentation on and targeted discovery into their metric-tracking dashboards. Defendants have not articulated any specific burden relating to News Plaintiffs' request for targeted discovery on their metric-tracking dashboards. Indeed, News Plaintiffs' request for documentation on Defendants' metric-tracking dashboards would make discovery into these metrics more efficient and avoid additional months of circular back and forth efforts. News Plaintiffs' request is akin to this Court's previous recognition that Defendants must provide guidance assisting News Plaintiffs' efficient discovery into unfamiliar data sources. *See* Ex. L (December 3, 2024 Discovery Hearing Transcript) at 77:16-17 (discussing Microsoft's output log data and suggesting that the parties work together so that News Plaintiffs can "learn how to use this dataset in the most efficient way"); 77:20-25 (recognizing that OpenAI training data is "a huge volume of data that the plaintiffs are not familiar with" and suggesting that the parties work together to "get [News Plaintiffs] up to speed so that they can be more efficient[.]").

Just as the Court ordered Defendants to provide News Plaintiffs with the information necessary to efficiently query Defendants' output logs and training data, Defendants should now be ordered to provide documentation on their metric-tracking dashboards so that News Plaintiffs can identify the specific metrics responsive to their requests and the parties can thereafter work together to develop a plan for Defendants to query these dashboards and produce the resulting data to News Plaintiffs.

---

[5] Microsoft and News Plaintiffs have a pending dispute about discovery into Copilot dated past October 1, 2024. Times Dkt. 487.

Respectfully,


Respectfully,


 _/s/ Ian Crosby_____
Ian B. Crosby
Susman Godfrey LLP

 _/s/ Steven Lieberman_____
Steven Lieberman
Rothwell Figg Ernst & Manbeck, P.C.


cc:      All Counsel of Record (via ECF)